**Texas Commission on Environmental Quality**
**Comments Regarding the United States Environmental Protection Agency**
**Draft Toxicological Review of Inorganic Arsenic: In Support of the Summary Information on the Integrated Risk Information System (IRIS)**
**Notice of Public Comment Period**
**75 FR 7477, February 19, 2010**
**Docket ID No. EPA–HQ–ORD–2010–0123**

The Texas Commission on Environmental Quality (TCEQ) provides the following comments on the United States Environmental Protection Agency (USEPA) announcement of the public comment period regarding its *Draft Toxicological Review of Inorganic Arsenic: In Support of the Summary Information on the Integrated Risk Information System (IRIS).*

On February 19, 2010, the USEPA published a Federal Register notice of a 60-day public comment period (ending April 20, 2010) for the *Draft Toxicological Review of Inorganic Arsenic: In Support of the Summary Information on the Integrated Risk Information System (IRIS)* (Federal Register/Vol. 75, No. 33/Friday, February 19, 2010/Notices). USEPA will only guarantee that comments submitted by March 26, 2010, will be provided to the Scientific Advisory Board in time for their meeting to consider the final draft EPA document. This final draft USEPA document (EPA/635/R-10/001) derives an oral slope factor (SFo) for arsenic to ultimately be published on IRIS. To the extent practicable in the time allotted by USEPA, Toxicology Division staff of the TCEQ have developed comments for USEPA consideration.

***60-Day Public Comment Period***

The 60-day comment period is insufficient for regulatory agencies and others to provide meaningful comments based on an in-depth review and analysis of the derivation of the final draft SFo. There is great complexity associated with multiple issues relevant to the assessment of arsenic risk due to oral exposure. The final draft document alone is 575 pages, with the Science Advisory Board (SAB) comments on three USEPA documents relevant to USEPA's final draft arsenic assessment being almost another 100 pages, and hundreds of pages (at a bare minimum) of other documents (e.g., National Research Council 1991 and 2001 reviews) and studies relevant to the assessment of risk due to oral arsenic exposure. Given the complexity and volume of relevant materials, it is impracticable for USEPA to expect detailed specific comments from external experts given the short period allowed for a critical review of the document and procedures employed by USEPA. To exacerbate the short review time problem, the 5-day Society of Toxicology 49$^{th}$ Annual Meeting (March 7-11) and the 3-day Alliance for Risk Assessment dose-response conference (*Beyond Science and Decisions: From Issue Identification to Dose-Response Assessment*, March 16-18) fall within the review period, and TCEQ staff and many other external expert peer reviewers will be in attendance. The 60-day comment period only allows a cursory review of the document at best, leads to a less-than-desirable level of peer review and transparency, and undermines confidence in the final draft SFo value. Consequently, TCEQ is only able to provide preliminary

comments on the final draft SFo value, barely scratching the surface of the document. The comment deadline should be extended at least 60 days past the current April 20th deadline to allow for a detailed review of the hundreds of pages of documents (at a bare minimum) and complex issues relevant to derivation of the final draft SFo for arsenic.

***Arsenic SFo***

The final draft SFo of 25.7 per mg/kg-day represents a 17-fold increase over the SFo currently on IRIS (1.5 per mg/kg-day). This is a significant change in the estimated carcinogenic potency of arsenic. Arsenic already has a relatively high SFo and such a large change would have far reaching regulatory implications. Thus, the final draft SFo deserves greater scrutiny than allowed by the 60-day public comment period. In addition to TCEQ's concerns, we understand both external groups and internal USEPA staff have expressed serious concerns about the final draft SFo. Brief discussions of four areas of TCEQ concern relevant to the toxicological basis for the derivation of the final draft SFo are provided below. This discussion is followed by comments on some practical implications that highlight the importance of EPA developing a scientifically-defensible SFo for arsenic.

***Toxicological Concerns with the Final Draft Arsenic SFo***

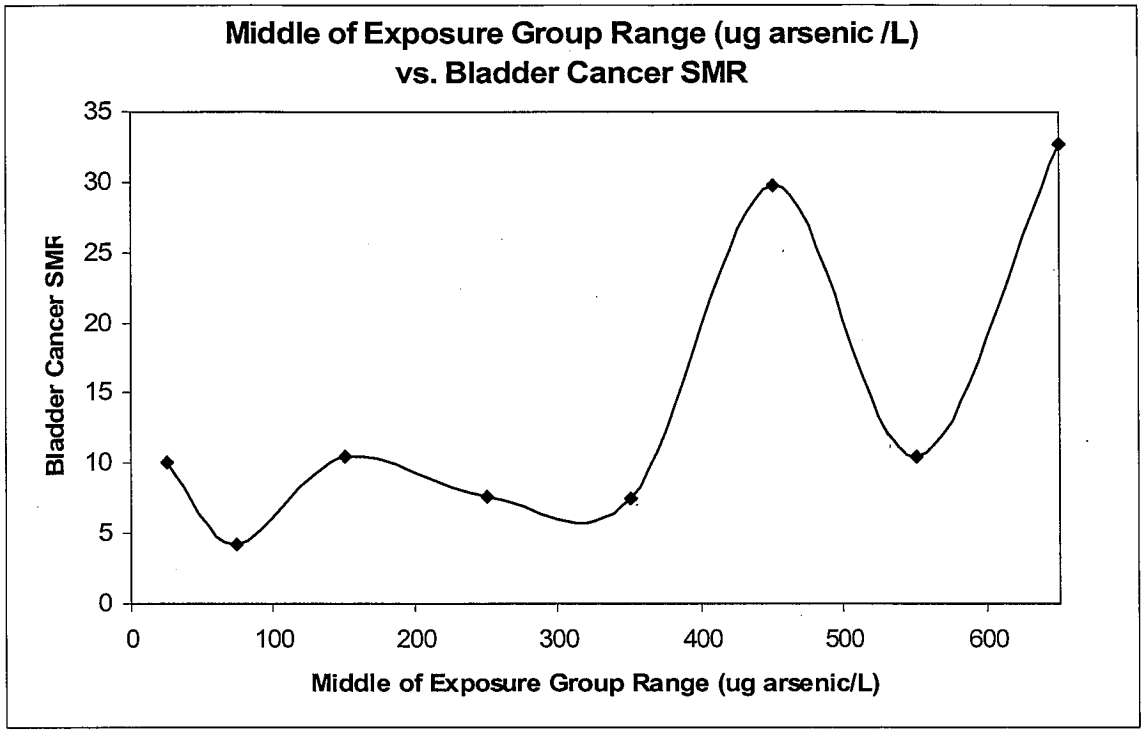
Water Intake and Non-water Arsenic Intake

The final draft USEPA document acknowledges that there is significant uncertainty associated with water intake (e.g., see page 117 "few precise data," "limited information"; page 120 "drinking water exposure information is not available for individual study subjects") and non-water arsenic intake (e.g., see page 118 "relatively little data," "considerable confusion about" how to include this; page 123 "data supporting this value are scarce") for the exposed Taiwanese populations. Data on variations in arsenic drinking water levels with time are also lacking. TCEQ believes it unreasonable to exclude arsenic intake from water used for cooking rice and produce (e.g., rice and yams are staples) from dietary intake for exposed populations for the primary analysis as indirect water intake estimates are around 1 L/day (page 124), significantly underestimating dose. Additionally, there are no drinking water arsenic data for the reference populations, and TCEQ has serious concerns about the reasonableness of assuming zero arsenic drinking water intake for reference populations. TCEQ also has significant concerns about assuming the same non-water arsenic intake (10 μg/day) for both the reference and exposed populations given that USEPA acknowledges that exposed populations may be exposed to 15-211 μg/day (50 μg/day average) from food (page 123). The examination of such issues in a sensitivity analysis does not alleviate USEPA's duty to derive the most accurate SFo possible in the primary analysis by incorporation of the most informed estimates possible for factors known to be critical for derivation of a reasonably predictive SFo (e.g., population-specific factors influencing total dose such as indirect water and food intake).

USEPA appears to lack data sufficient to establish the extent to which total arsenic exposure (i.e., dose) differed for the exposed and "nonexposed" populations, making derivation of a reasonably accurate SFo problematic. Accurate water intake and non-water arsenic intake data are critical in deriving defensible dose estimates and a scientifically-defensible carcinogenic assessment, but are lacking. The admitted absence of accurate dose estimates due to lack of good water intake and non-water arsenic intake data precludes the conduction/derivation of an accurate dose-response assessment and SFo.

Dose-Response Data

USEPA used lung and bladder mortality data from Morales et al. (2000) for the dose-response assessment for the final draft SFo. Morales et al. (2000) uses these mortality data to calculate standardized mortality ratios (SMRs) and notes, "Although the computed SMRs display a large amount of noise, there appear to be higher SMRs at high exposure levels compared to exposures in the lower range, especially for bladder and lung cancer." To say that there is "noise" in the SMRs over the eight exposure categories is an understatement. Dose-response is the cornerstone of toxicology, but the lung and bladder mortality data (SMRs) from Morales et al. (2000) provide a poor basis for dose-response assessment as a dose-response is not apparent and not monotonic. Breaking the data down into the form of age-specific person-years at risk and cancer deaths does not improve the basis for dose-response assessment; it only obscures the lack of a good dose-response which is readily apparent from examination of the SMRs. For example, for lung cancer, SMRs greater than 3 were essentially only obtained for drinking water levels greater than 400 μg/L, which does not indicate a particularly strong dose-response. Even at 500-600 μg/L, the SMR was only 3.32. For bladder cancer, the dose-response data from Morales et al. (2000) and used by USEPA do a poor job of characterizing the shape of the dose-response curve, as can be seen from the figure below (line added for emphasis).




The cancer guidelines (USEPA 2005) recommend use of enough dose groups to provide an indication of the shape of the dose-response curve, as characterization of the shape of the dose-response curve is important in providing relevant dose-response data for assessing human risk. A relatively large exposure range should make it relatively easy to discern the shape of any underlying dose-response curve in a well-conducted study. However, despite the eight exposure groups in Morales et al. (2000), the figure above illustrates that the shape of the dose-response curve for bladder cancer, which had the highest SMRs by far, has not been adequately defined by the dose-response data selected by USEPA for derivation of the SFo. As an example, the SMR for the 0-50 μg/L exposure group (plotted at 25 μg/L) is higher than that for the 300-400 μg/L exposure group (plotted at 350 μg/L), and similar to that for the 500-600 μg/L exposure group (plotted at 550 μg/L). The ability to fit a line through data points does not necessarily mean that the underlying data adequately define the shape of the dose-response curve, including the critical low dose region. Based on the above considerations, the underlying data modeled by USEPA provide a poor basis for dose-response assessment.

Biological Effects of Ionizing Radiations (BEIR) IV Model

Appendix E to the final draft USEPA document indicates that a modified BEIR IV model was used, which takes as inputs the dose-response "b" coefficient, background cancer incidence data, and age-specific mortality data, to estimate bladder and lung cancer incidence for the US population. A modification by Gail et al. (1999) was used to obtain estimates of incidence within multi-year age strata, which itself would have associated uncertainty. The short time allotted for review is inadequate for a full examination of the appropriateness of the modified BEIR IV methodology used by USEPA (and a plethora

of other potential issues). However, generally, the BEIR IV methodology for calculating excess risk is mathematically correct only when the specified response is mortality and mortality rates are used, not when the specified response is mortality and incidence rates are used, or when the specified response is incidence and incidence rates are used with BEIR IV equations which have not been appropriately derived for incidence. The beta or "b" value used by USEPA for *incidence* calculations at a given dose is based on *mortality* (pages 127, E-1), which is inappropriate. Additionally, BEIR IV equations are for mortality and may not be used for incidence without modification (i.e., derivation of appropriate BEIR IV equations specifically for incidence). This potential error is demonstrated in Appendix 1 to these comments. Although time did not allow for a more detailed review, USEPA does not indicate that any specific alterations were made to BEIR IV equations to account for incidence as the response. Therefore, TCEQ believes that USEPA may have used inappropriate BEIR IV methodology.

***Some Practical Implications of Final Draft Arsenic SFo***

USEPA's Soil Screening Levels

The current USEPA regional screening level (RSL) for inorganic arsenic in residential soil is 0.39 mg/kg. The US Geological Survey reports the mean for arsenic in soil is 7.2 mg/kg (ATSDR 2007), and TCEQ uses a median background arsenic concentration for Texas soils of 5.9 mg/kg. Thus, the current residential soil RSL is already 18 times less than typical background soil arsenic concentrations. Adoption of the final draft SFo would reduce the current USEPA residential soil RSL by a factor of 17 to approximately 0.02 mg/kg at a conservative target excess risk level of 1 in 1,000,000. Even a residential soil RSL of 2 mg/kg corresponding to the upper end of the USEPA acceptable risk range (1 in 10,000) using the final draft SFo would be below typical background concentrations, making achievement of acceptable risk as defined by USEPA practically impossible at remediation sites. More importantly, this analysis would imply that typical naturally-occurring levels of arsenic in residential soil are unsafe for human contact.

In regard to individual excess lifetime cancer risk (IELCR), USEPA states on their website (http://www.epa.gov/oust/rbdm/sctrlsgw.htm)), "The IELCR represents the incremental (over background) probability of an exposed individual's getting cancer (i.e., a risk occurring in excess of or above and beyond other risks for cancer such as diet, smoking, heredity). Cleanup standards calculated on the basis of excess risk limits correspond to *allowable levels in excess of the background concentrations of the chemicals of concern normally present in the source media*" (emphasis added). Since regulatory agencies are concerned with regulating excess risk (i.e., risk over natural background), the risk due to naturally-occurring background soil arsenic levels should be excluded from comparisons to the USEPA acceptable risk range. In effect, this is typically accomplished by USEPA acknowledging that although above the RSL or proposed remediation goal (PRG), soil arsenic levels at a remediation site are within background so no action is necessary in regard to arsenic. In a more strict sense, however, since per USEPA regulatory agencies calculate cleanup values based on excess risk over background, the soil RSL/PRG could be added to a representative background

concentration to derive a comparison value which represents a regulatory acceptable level of excess risk (i.e., risk over background).

Implications for Food and Drinking Water Safety: Typical Dietary Exposure, Rice Consumption, Drinking Water, and Fish/Shellfish Consumption as Examples

A scientifically-defensible and realistic dose-response assessment for inorganic arsenic is critical given the grave implications of the final draft SFo for the US food and water supply. The examples below illustrate how estimates of risk due to dietary exposure to inorganic arsenic using the final draft SFo may have dire consequences on the perceived safety of US food and drinking water.

*Typical Dietary Exposure*

Using the final draft SFo for inorganic arsenic results in excess cancer risk estimates from dietary exposure exceeding the USEPA acceptable risk range (1 in 1,000,000 to 1 in 10,000). ATSDR (2007) reports the mean average US adult intake of inorganic arsenic is around 10.22 µg/day (range of 0.93-104.89 µg/day) based on a study (MacIntosh et al. 1997) which utilized residue data collected for the Food & Drug Administration Total Diet Study. Using the final draft SFo, excess calculated cancer risk would range from about 3.4 in 10,000 to 3.9 in 100, with an average calculated risk of about 3.8 in 1,000 due to dietary exposure. Even the calculated risk associated with the lower end of dietary inorganic arsenic exposure (3.4 in 10,000) would exceed the upper end of the USEPA acceptable risk range (1 in 10,000), and the calculated risk associated with the high end of dietary exposure would be 390 times the upper end of acceptable risk. Such analyses would imply that the US diet results in arsenic risk that is considered unsafe from a regulatory perspective.

*Rice Consumption*

In regard to eating rice specifically, the average excess risk for US adult (70 kg) rice eaters would be calculated at around 1.7 in 1,000 based on an average intake of 61.2 g dry rice/day (around 1 cup cooked) based on National Health & Nutrition Examination Survey data (Batres-Marquez and Jensen 2005) with 0.276 µg total arsenic/g US white rice and 27% of the total arsenic as inorganic arsenic (Williams et al. 2005). Even using a US adult average for rice intake that includes non-rice eaters (11.4 g dry rice/day) would still result in an excess risk of 3.1 in 10,000 for white rice, which exceeds the upper end (1 in 10,000) of USEPA's risk management range. Risk estimates would be higher for US brown rice than white rice due to a higher average percentage of total arsenic being inorganic (51%) (Williams et al. 2005), with average excess risk for US adult rice eaters being around 2.6 in 1,000 (26 times higher than the upper end of USEPA's risk management range). Such analyses would imply that rice and other food items (e.g., fish/shellfish) are unsafe for human. Consequently, there may be a potential for unwarranted advisories or warning labels on certain foods.

*Drinking Water*

Another implication of the draft final SFo is that the water used to prepare the rice (see example above) is itself by this calculation unsafe for human consumption. Drinking water in the US generally contains an average of 2 μg/L of arsenic (ATSDR 2007). Based on final draft SFo estimates, USEPA indicates that drinking water concentrations corresponding to 1 in 10,000 combined cancer risks for males and females are 0.21 and 0.14 μg/L, respectively. The implication is that on average all across the US, people's drinking water contains arsenic levels that exceed the upper end of the USEPA acceptable risk range (1 in 10,000) by approximately 10-14 times. In other words, on average, the level of arsenic in the nation's drinking water supply is unsafe.

For bladder cancer alone, the incidence risk calculated by USEPA based on final draft values for males/females is 3.1E-04 per μg/L. Therefore, based on 2 μg/L as an average drinking water concentration, the estimated bladder cancer risk for the US population would be 6.2 per 10,000 or 62 per 100,000. However, the actual occurrence of bladder cancer in the US is about 23 cases per 100,000 (males/females combined). It would take 3 times the actual bladder cancer incidence for US males/females combined to even make possible the 62 cases per 100,000 estimated due to arsenic exposure from drinking water alone. Thus, the incidence risk calculated by USEPA final draft values for bladder cancer appears to be inaccurate and overly conservative. Proceeding with this SFo will unnecessarily alarm the public by giving a greater perception of harm and risk than is actually taking place.

*Fish/Shellfish Consumption*

Shellfish and other marine foods contain the highest arsenic concentrations and are the largest dietary source of arsenic. Based on an FDA Total Diet Study, ATSDR (2007) reports that concentrations in canned tuna, fish sticks, haddock, and boiled shrimp were 0.609-1.470, 0.380-2.792, 0.510-10.430, and 0.290-2.681 mg/kg dry weight, respectively. The foods with the highest mean arsenic levels were haddock, canned tuna, fish sticks, shrimp, and fish sandwiches, with arsenic concentrations ranging from 0.568-5.33 mg/kg dry weight. Most recent studies show an arsenic concentration range of 0.82-37 mg/kg dry weight for fish (e.g., flounder, cod, sole, tuna), mussels, clams, oysters, shrimp, and blue crab, including fish, blue crabs, shrimp, mussels, and oysters from Texas (0.82-9.67 mg/kg) (see Galveston Bay/Gulf of Mexico results in Table 6-4 of ATSDR 2007).

The general consensus in the literature is that approximately 10% of the arsenic in the edible parts of marine fish and shellfish is inorganic arsenic (ATSDR 2007). A 10% adjustment to these reported arsenic levels in fish yields an inorganic arsenic concentration range of 0.029-3.7 mg inorganic arsenic/kg dry weight. Using the final draft SFo, a saltwater fish ingestion rate of 15 g/day (only two fish meals per month approximately), and an adult body weight (70 kg), the fish tissue concentration corresponding to the upper end of the USEPA acceptable risk range (1 in 10,000) is 0.017 mg inorganic arsenic/kg dry weight. The range of estimated inorganic arsenic levels in all these fish/seafood items (0.029-3.7 mg inorganic arsenic/kg) exceeds the fish tissue

concentration calculated at the upper end of acceptable excess risk (1 in 10,000) using the final draft SFo. Regarding Texas specifically, the range of estimated inorganic arsenic levels in Galveston Bay/Gulf of Mexico seafood (0.082-0.967 mg/kg dry weight based on Table 6-4 in ATSDR 2007) is 5-57 times higher than the fish tissue concentration (0.017 mg/kg) calculated at the upper end of acceptable excess risk using the final draft SFo. These analyses would imply that fish/shellfish in the US diet are unsafe for human consumption from a regulatory perspective. In turn, a determination of unacceptable risk due to arsenic in fish tissue would likely cause more waterbodies to be listed as impaired unnecessarily. As a result, there could be future inappropriate regulatory actions and unneeded expenditure of resources to investigate and try to reduce arsenic. There could also be negative public health consequences from such impairments, because fish consumption and the associated health benefits would decrease due to the false perception that arsenic is making fish unsafe to eat.

***References***

Agency for Toxic Substances and Disease Registry (ATSDR). 2007. Toxicological Profile for Arsenic.

Batres-Marquez SP, Jensen HH. 2005. Rice consumption in the United States: new evidence from food consumption surveys. Staff Report 05-SR 100. Center for Agricultural and Rural Development, Iowa State University, Ames, Iowa.

Gail MH, Kessler M, Midthune D, et al. 1999. Two approaches for estimating disease prevalence from population-based registries of incidence and total mortality. Biometrics 55:1137–1144.

MacIntosh DL, Williams PL, Hunter DJ, et al. 1997. Evaluation of a food frequency questionnaire-food composition approach for estimating dietary intake of inorganic arsenic and methylmercury. Cancer Epidemiol Biomarkers Prev 6(12):1043-1050.

Morales KH, Ryan L, Kuo T-L, et al. 2000. Risk of internal cancers from arsenic in drinking water. Environ Health Perspect 108:655–661.

Williams PN, Price AH, Raab A, et al. 2005. Variation in arsenic speciation and concentration in paddy rice related to dietary exposure. Environ Sci Technol 39:5531-5540.

## APPENDIX 1:
## CALCULATING EXCESS RISK WHEN SPECIFIED RESPONSE IS MORTALITY VERSUS INCIDENCE

**Issues in Quantitative Epidemiology**
**Calculating Excess Risk When Specified Response is Mortality**
**Vs When the Specified Response is Incidence**

Robert L. Sielken Jr., Ph.D.
Ciriaco Valdez-Flores, Ph.D., P.E.
Sielken & Associates Consulting, Inc.
3833 Texas Avenue, Suite, 230, Bryan, TX 77802
Tel: 979-846-5175; Fax: 979-846-2671; Email: SielkenAssoc@aol.com

**TCEQ Contract 582-7-81521**

The BEIR IV methodology for calculating excess risk is mathematically correct when the specified response is mortality; however, the BEIR IV methodology is mathematically incorrect when the specified response is incidence (not death).

The following slides are divided into two presentations. The first presentation provides a step-by-step derivation of the BEIR IV methodology when the specified response is mortality. This presentation directly parallels the same derivation in BEIR IV. The second presentation provides a step-by-step derivation that is "parallel" to that in the first presentation except that in the second presentation the specified response is incidence (not death). However, the steps and result are fundamentally different when the specified response is incidence (not death) than when the response is death.

The fact that the "result" (i.e., the mathematical formula for calculating excess risk) is different when the response is mortality than it is when the response is incidence means that when the response is incidence (not death) the excess risk cannot be validly calculated using the formula (BEIR IV methodology) for death.

## The First Presentation: Issues in Quantitative Epidemiology: Calculating Excess Risk: When Specified Response is Mortality

***Calculating Excess Risk using Actuarial Method or Life Table Method.*** This way of calculating excess risks from a RR function is the implementation of the methodology described in "BEIR IV. Health Risks of Radon and Other Internally Deposited Alpha-Emitters. Committee on the Biological Effects of Ionizing Radiations. Board on Radiation Effects Research Commission of Life Sciences. National Research Council. National Academy Press, Washington, DC, 1988."

BEIR IV:

Derivation of Formulas:
(Using notation in BEIR report)

$i = 1, 2, ..., T$

i = index for the years for a person's life

year i is the year from the person's (i-1)-th birthday
to his (or her) i-th birthday

i=1 refers to the year from birth to the 1st birthday
i=1 = age 0
...
i=7 refers to the year from the 6-th birthday to the 7-th birthday
i=7 = age 6

BEIR IV: Derivation of Formulas:

$i = 1, 2, ..., T$

q(i) = probability of surviving year i
when all causes of death are acting
conditional on the person surviving through year i -1

q(7) = probability of reaching a person 's 7-th birthday
given that he reached his 6 -th birthday

$q(7) = P( \text{Death} \geq 7 \mid \text{Death} \geq 6 )$

h(i)* = mortality rate due to all causes in year i
conditional on the person surviving through year i -1

$q(i) = \exp[ - h(i)^* ]$

1 - q(i) = probability of death in year i
conditional on the person surviving through year i -1

BEIR IV: Derivation of Formulas:

$i = 1, 2, ..., T$

$q(i)$ = probability of surviving year i
when all causes of death are acting
conditional on the person surviving through year i-1

$h(i)^*$ = mortality rate due to all causes in year i
conditional on the person surviving through year i-1

$q(i) = \exp[ - h(i)^* ]$

$1 - q(i)$ = probability of death in year i
conditional on the person surviving through year i-1

$h(i)$ = response (e.g., lung cancer) mortality rate in year i
conditional on the person not having the response
through year i-1

BEIR IV: Derivation of Formulas:

$i = 1, 2, ..., T$

$q(i)$ = probability of surviving year i
when all causes of death are acting
conditional on the person surviving through year i-1

$h(i)^*$ = mortality rate due to all causes in year i
conditional on the person surviving through year i-1

$h(i)$ = response (e.g., lung cancer) mortality rate in year i
conditional on the person not having the response
through year i-1

$S(1,i)$ = probability of surviving up to year i is the product of surviving each prior year:

$$S(1,i) = q(1) \times q(2) \times ... \times q(i-1) \quad \text{with } S(1,1) = 1.0.$$

$S(1,i) \times [ 1 - q(i) ]$ = probability of surviving up to year i and then dying (from any cause) in year i

## BEIR IV: Derivation of Formulas:

$i = 1, 2, ..., T$

$q(i)$ = probability of surviving year i
when all causes of death are acting
conditional on the person surviving through year i-1

$h(i)^*$ = mortality rate due to all causes in year i
conditional on the person surviving through year i-1

$h(i)$ = response (e.g., lung cancer) mortality rate in year i
conditional on the person not having the response
through year i-1

$S(1,i)$ = probability of surviving up to year i

$S(1,i) \times [\, 1 - q(i) \,]$ = probability of surviving up to year i and
then dying (from any cause) in year i

$h(i)/h(i)^*$ = proportion of deaths in year i due to the response

$[\, h(i)/h(i)^* \,] \times S(1,i) \times [1 - q(i) \,]$ = probability of surviving i-1 years
and dying of response in year i

## BEIR IV: Derivation of Formulas:

$i = 1, 2, ..., T$

$q(i)$ = probability of surviving year i
when all causes of death are acting
conditional on the person surviving through year i-1

$h(i)^*$ = mortality rate due to all causes in year i
conditional on the person surviving through year i-1

$h(i)$ = response (e.g., lung cancer) mortality rate in year i
conditional on the person not having the response
through year i-1

$S(1,i)$ = probability of surviving up to year i = $q(1) \times q(2) \times ... \times q(i-1)$

$S(1,i) \times [\, 1 - q(i) \,]$ = probability of surviving up to year i and
then dying (from any cause) in year i

$h(i)/h(i)^*$ = proportion of deaths in year i due to the response

$[\, h(i)/h(i)^* \,] \times S(1,i) \times [1 - q(i) \,]$ = probability of surviving i-1 years
and dying of response in year i

$R_0 = \sum_{i=1,...,T} [\, h(i)/h(i)^* \,] \times S(1,i) \times [1 - q(i) \,]$
= probability of a response mortality in the first T years of life
(i.e., up to the T-th birthday, age T ) at dose 0
(no exposure in addition to background exposure)

## BEIR IV: Derivation of Formulas: Risk with exposure

i=1, 2, ..., T

q(i) = probability of surviving year i without exposure
when all causes of death are acting
conditional on the person surviving through year i-1

h(i) = response (e.g., lung cancer) mortality rate in year i without exposure
conditional on the person not having the response through year i-1

h(i)* = mortality rate due to all causes in year i without exposure
conditional on the person surviving through year i-1

f(i) = proportional effect (multipler) in year i assuming a proportional hazards model for the effect of exposure of the form $h(i) \times f(i)$
$f(i) = [\ 1 + e(i)\ ]$ if the multiplier is a linear function

$h(i) \times f(i)$ = response (e.g., lung cancer) mortality rate in year i with exposure conditional on the person not having the response through year i-1

$h(i) \times [\ f(i) - 1\ ]$ = increase in response mortality rate in year due to exposure

## BEIR IV: Derivation of Formulas: Risk with exposure

i=1, 2, ..., T

q(i) = probability of surviving year i without exposure
when all causes of death are acting
conditional on the person surviving through year i-1

h(i) = response (e.g., lung cancer) mortality rate in year i without exposure
conditional on the person not having the response through year i-1

h(i)* = mortality rate due to all causes in year i without exposure
conditional on the person surviving through year i-1

f(i) = proportional effect (multipler) in year i assuming a proportional hazards model for the effect of exposure of the form $h(i) \times f(i)$
$f(i) = [\ 1 + e(i)\ ]$ if multiplier is a linear function

$h(i) \times f(i)$ = response (e.g., lung cancer) mortality rate in year i with exposure conditional on the person not having the response through year i-1

$h(i) \times [\ f(i) - 1\ ]$ = increase in response mortality rate in year due to exposure

$h(i)^* + h(i) \times [\ f(i) - 1\ ]$ = mortality rate due to all causes in year i with exposure conditional on the person surviving through year i-1

## BEIR IV: Derivation of Formulas: Risk with exposure

i=1, 2, ..., T

q(i) = probability of surviving year i without exposure
when all causes of death are acting
conditional on the person surviving through year i-1

h(i) = response (e.g., lung cancer) mortality rate in year i without exposure
conditional on the person not having the response through year i-1

h(i)* = mortality rate due to all causes in year i without exposure
conditional on the person surviving through year i-1

f(i) = proportional effect (multipler) in year i assuming a proportional hazards
model for the effect of exposure of the form h(i) × f(i)
f(i) = [ 1 + e(i) ] if multiplier is a linear function

h(i) × f(i) = response (e.g., lung cancer) mortality rate in year i with exposure
conditional on the person not having the response
through year i-1

h(i) × [ f(i) - 1 ] = increase in response mortality rate in year i due to exposure

h(i)* + h(i) × [ f(i) - 1 ] = mortality rate due to all causes in year i with exposure
conditional on the person surviving through year i-1

$\exp \{ - h(i)^* - h(i) \times [ f(i) - 1 ] \}$ = probability with exposure of surviving year i conditional on person surviving thru year i-1

$q(i) \times \exp \{ - h(i) \times [ f(i) - 1 ] \}$ = probability with exposure of surviving year i conditional on person surviving thru year i-1

## BEIR IV: Derivation of Formulas: Risk with exposure

q(i) = probability of surviving year i without exposure
when all causes of death are acting conditional on the person surviving through year i-1

h(i) = response (e.g., lung cancer) mortality rate in year i without exposure
conditional on the person not having the response through year i-1

h(i)* = mortality rate due to all causes in year i without exposure
conditional on the person surviving through year i-1

f(i) = proportional effect (multipler) in year i assuming a proportional hazards
model for the effect of exposure of the form h(i) × f(i); f(i) = [ 1 + e(i) ] if multiplier is a linear function

h(i) × f(i) = response (e.g., lung cancer) mortality rate in year i with exposure
conditional on the person not having the response through year i-1

h(i) × [ f(i) - 1 ] = increase in response mortality rate in year due to exposure

h(i)* + h(i) × [ f(i) - 1 ] = mortality rate due to all causes in year i with exposure
conditional on the person surviving through year i-1

$\exp \{ - h(i)^* - h(i) \times [ f(i) - 1 ] \}$ = probability with exposure of surviving year i conditional on person surviving thru year i-1

$q(i) \times \exp \{ - h(i) \times [ f(i) - 1 ] \}$ = probability with exposure of surviving year i conditional on person surviving thru year i-1

$$q(1) \times \exp \{ - h(1) \times [ f(1) - 1 ] \} \times \ldots \times q(i-1) \times \exp \{ - h(i-1) \times [ f(i-1) - 1 ] \}$$
$$= S(1,i) \times \exp \left( - \sum_{k=1,\ldots,i-1} \{ - h(k) \times [ f(k) - 1 ] \} \right)$$
= probability of surviving up to year i with exposure

$$S(1,i) \times \exp \left( - \sum_{k=1,\ldots,i-1} \{ -h(k) \times [ f(k) - 1 ] \} \right) \times (1 - q(i) \times \exp\{ -h(i) \times [f(i) - 1] \} )$$
= probability with exposure of surviving up to year i and then dying (from any cause) in year i

## BEIR IV: Derivation of Formulas: Risk with exposure

**q(i) = probability of surviving year i without exposure when all causes of death are acting conditional on the person surviving through year i-1**

**h(i) = response (e.g., lung cancer) mortality rate in year i without exposure conditional on the person not having the response through year i-1**

**h(i)* = mortality rate due to all causes in year i without exposure conditional on the person surviving through year i-1**

**f(i) = proportional effect (multipler) in year i assuming a proportional hazards model for the effect of exposure of the form h(i) × f(i); f(i) = [ 1 + e(i) ] if multiplier is a linear function**

**h(i) × f(i) = response (e.g., lung cancer) mortality rate in year i with exposure conditional on the person not having the response through year i-1**

**h(i) × [ f(i) - 1 ] = increase in response mortality rate in year due to exposure**

**h(i)* + h(i) × [ f(i) - 1 ] = mortality rate due to all causes in year i with exposure conditional on the person surviving through year i-1**

**exp { -h(i)* - h(i) × [ f(i) - 1 ] } = probability with exposure of surviving year i conditional on person surviving thru year i-1**

**q(i) × exp { - h(i) × [ f(i) - 1 ] } = probability with exposure of surviving year i conditional on person surviving thru year i-1**

**q(1) × exp { - h(1) × [ f(1) - 1 ] } × ... × q(i -1) × exp { - h(i -1) × [ f(i -1) - 1 ] }**
**= S(1,i) × exp ( - $\sum_{k=1,...,i-1}$ { - h(k) × [ f(k) - 1 ] } )= probability of surviving up to year i with exposure**

**S(1,i) × exp ( - $\sum_{k=1,...,i-1}$ {-h(k)×[ f(k) - 1 ] } ) × (1 - q(i)×exp{-h(i)×[f(i) - 1] } )**
**= probability with exposure of surviving up to year i and then dying (from any cause) in year i**
**{ h(i) × f(i) } / { h(i)* + h(i) × [ f(i) - 1] }**
**= proportion of deaths in year i due to the response with exposure**
**({h(i)×f(i)}/{h(i)*+h(i)×[f(i)-1]}) × S(1,i) × exp(- $\sum_{k=1,...,i-1}${-h(k)×[f(k)-1]}) × (1-q(i)×exp{-h(i)×[f(i)-1] } )**
**= probability of surviving i-1 years and dying of response in year i with exposure**

## BEIR IV: Derivation of Formulas: Risk with exposure

( {h(i)×f(i)}/{h(i)*+h(i)×[f(i)-1]} )
× S(1,i) × exp(- $\sum_{k=1,...,i-1}${-h(k)×[f(k)-1]})
× (1-q(i)×exp{-h(i)×[f(i)-1] } )
= probability of surviving i-1 years
and dying of response in year i with exposure

$R_{exposure} = \sum_{i=1,...,T}$
( {h(i) × f(i)} / {h(i)* + h(i) × [ f(i)-1 ] } )
× S(1,i) × exp(- $\sum_{k=1,...,i-1}$ { -h(k) × [ f(k)-1 ] } )
× (1 - q(i) × exp { - h(i) × [ f(i) - 1 ] } )

= probability of a response mortality in the first T years of life (i.e., up to the T-th birthday, age T ) with exposure (with exposure in addition to the background exposure)

## BEIR IV: Risks

$R_0 = \sum_{i=1,\ldots,T} [\, h(i)/h(i)^* \,] \times S(1,i) \times [1 - q(i)\,]$

= probability of a response mortality in the first T years of life (i.e., up to the T-th birthday, age T ) at dose 0

(no exposure in addition to background exposure)

$$R_{exposure} = \sum_{i=1,\ldots,T} ( \{h(i) \times f(i)\} / \{h(i)^* + h(i) \times [\, f(i)-1 \,]\} ) \times S(1,i) \times \exp(- \sum_{k=1,\ldots,i-1} \{ -h(k) \times [\, f(k)-1 \,]\} ) \times (1 - q(i) \times \exp \{ - h(i) \times [\, f(i) - 1 \,]\} )$$

= probability of a response mortality in the first T years of life (i.e., up to the T-th birthday, age T ) with exposure (with exposure in addition to the background exposure)

Added Risk = $R_{exposure} - R_0$

Extra Risk = $( R_{exposure} - R_0 ) / ( 1 - R_0 )$

Excess Risk = either Added Risk or Extra Risk

## The Second Presentation: 3.1 Issues in Quantitative Epidemiology: Calculating Excess Risk: When Specified Response is Incidence

Calculating Excess Risk using Actuarial Method or Life Table Method. The following derivation for the situation in which the specified response is incidence (not death) "parallels" the derivation in BEIR IV; however, the derivation and result are necessarily different for incidence than for mortality.

"BEIR IV. Health Risks of Radon and Other Internally Deposited Alpha-Emitters. Committee on the Biological Effects of Ionizing Radiations. Board on Radiation Effects Research Commission of Life Sciences. National Research Council. National Academy Press, Washington, DC, 1988."

Derivation of Formulas:
(Using notation in BEIR report)

i = 1, 2, ..., T

i = index for the years for a person's life

year i is the year from the person's (i-1)-th birthday
to his (or her) i-th birthday

i=1 refers to the year from birth to the 1st birthday
i=1 = age 0
...
i=7 refers to the year from the 6-th birthday to the 7-th birthday
i=7 = age 6

Derivation of Formulas:

i = 1, 2, ..., T

q(i) = probability of surviving year i
when all causes of death are acting
conditional on the person surviving through year i-1

q(7) = probability of reaching a person's 7-th birthday
given that he reached his 6-th birthday

$q(7) = P( \text{Death} \geq 7 \mid \text{Death} \geq 6 )$

h(i)* = mortality rate due to all causes in year i
conditional on the person surviving through year i-1

q(i) = exp[ - h(i)* ] -- definition of hazard rate

1 - q(i) = probability of death in year i
conditional on the person surviving through year i-1

Derivation of Formulas:

i = 1, 2, ..., T

q(i) = probability of surviving year i
when all causes of death are acting
conditional on the person surviving through year i-1

h(i)* = mortality rate due to all causes in year i
conditional on the person surviving through year i-1

q(i) = exp[ - h(i)* ]

1 - q(i) = probability of death in year i
conditional on the person surviving through year i-1

h(i) = response (e.g., lung cancer) incidence rate in year i
conditional on the person not having the response
through year i-1

Note that h(i) is NOT part of h(i)*,
because h(i) refers to incidence and h(i)* refers to death.

Derivation of Formulas:

i = 1, 2, ..., T

q(i) = probability of surviving year i
when all causes of death are acting
conditional on the person surviving through year i-1

h(i)* = mortality rate due to all causes in year i
conditional on the person surviving through year i-1

q(i) = exp[ - h(i)* ]

1 - q(i) = probability of death in year i
conditional on the person surviving through year i-1

h(i) = response (e.g., lung cancer) incidence rate in year i
conditional on the person not having the response
through year i-1

qr(i) = exp[ - h(i) ] = probability of no response in year i
conditional on the person not responding through year i-1

1 -qr(i) = probability of response (incidence) in year i
conditional on the person not responding through year i-1

## Derivation of Formulas:

i = 1, 2, ..., T

q(i) = probability of surviving year i
when all causes of death are acting
conditional on the person surviving through year i-1

h(i)* = mortality rate due to all causes in year i
conditional on the person surviving through year i-1

h(i) = response (e.g., lung cancer) incidence rate in year i
conditional on the person not having the response
through year i-1

qr(i) =probability of no response (incidence) in year i
conditional on the person not responding through year i-1

S(1,i) = probability of surviving up to year i is the product of surviving each prior year:

$$S(1,i) = q(1) \times q(2) \times \ldots \times q(i-1) \quad \text{with } S(1,1) = 1.0.$$

SR(1,i) = probability of no response up to year i is the product of no response in each prior year:

$$SR(1,i) = qr(1) \times qr(2) \times \ldots \times qr(i-1) \quad \text{with } SR(1,1) = 1.0.$$

## Derivation of Formulas:

**i = 1, 2, ..., T**

**q(i) = probability of surviving year i when all causes of death are acting conditional on the person surviving through year i-1**

**h(i)* = mortality rate due to all causes in year i conditional on the person surviving through year i-1**

**h(i) = response (e.g., lung cancer) incidence rate in year i conditional on the person not having the response through year i-1**

**qr(i) =probability of no response (incidence) in year i conditional on the person not responding through year i-1**

**S(1,i) = probability of surviving up to year i is the product of surviving each prior year:**

$$S(1,i) = q(1) \times q(2) \times \ldots \times q(i-1) \quad \text{with } S(1,1) = 1.0.$$

**SR(1,i) = probability of no response up to year i is the product of no response in each prior year:**

$$SR(1,i) = qr(1) \times qr(2) \times \ldots \times qr(i-1) \quad \text{with } SR(1,1) = 1.0.$$

$S(1,i) \times SR(1,i) \times [1 - q(i) \times qr(i)]$ = probability of surviving to year i, not responding before year i, and then dying (from any cause) or having the response in year i

## Derivation of Formulas:

i = 1, 2, ..., T

q(i) = probability of surviving year i when all causes of death are acting conditional on the person surviving through year i-1

h(i)* = mortality rate due to all causes in year i conditional on the person surviving through year i-1

h(i) = response (e.g., lung cancer) incidence rate in year i conditional on the person not having the response through year i-1

qr(i) =probability of no response (incidence) in year i conditional on the person not responding through year i-1

S(1,i) = probability of surviving up to year i is the product of surviving each prior year:
$S(1,i) = q(1) \times q(2) \times ... \times q(i-1)$ with $S(1,1) = 1.0$.

SR(1,i) = probability of no response up to year i is the product of no response in each prior year:
$SR(1,i) = qr(1) \times qr(2) \times ... \times qr(i-1)$ with $SR(1,1) = 1.0$.

$S(1,i) \times SR(1,i) \times [1 - q(i) \times qr(i)]$ = probability of surviving to year i, not responding before year i, and then dying (from any cause) or having the response in year i

A person is "observed" in year i if that person either dies in year i or has the specified response (incidence) in year i.

$h(i)/ [ h(i)^* + h(i) ]$ = proportion of observations (deaths plus incidences) in year i due to the response

$\{ h(i)/ [ h(i)^* + h(i) ] \} \times S(1,i) \times SR(1,i) \times [1 - q(i) \times qr(i) ]$ = probability of surviving to year i, not responding before year i, and then having the response (incidence) in year i

## Derivation of Formulas:

i = 1, 2, ..., T

q(i) = probability of surviving year i when all causes of death are acting conditional on the person surviving through year i-1

h(i)* = mortality rate due to all causes in year i conditional on the person surviving through year i-1

h(i) = response (e.g., lung cancer) incidence rate in year i conditional on the person not having the response through year i-1

qr(i) =probability of no response (incidence) in year i conditional on the person not responding through year i-1

S(1,i) = probability of surviving up to year i is the product of surviving each prior year:
$S(1,i) = q(1) \times q(2) \times ... \times q(i-1)$ with $S(1,1) = 1.0$.

SR(1,i) = probability of no response up to year i is the product of no response in each prior year:
$SR(1,i) = qr(1) \times qr(2) \times ... \times qr(i-1)$ with $SR(1,1) = 1.0$.

$S(1,i) \times SR(1,i) \times [1 - q(i) \times qr(i)]$ = probability of surviving to year i, not responding before year i, and then dying (from any cause) or having the response in year i

$h(i)/ [ h(i)^* + h(i) ]$ = proportion of observations (deaths plus incidences) in year i due to the response

$\{ h(i)/ [ h(i)^* + h(i) ] \} \times S(1,i) \times SR(1,i) \times [1 - q(i) \times qr(i) ]$ = probability of surviving to year i, not responding before year i, and then having the response (incidence) in year i

$R_0 = \sum_{i=1,...,T} \{ h(i) / [ h(i)^* + h(i) ] \} \times S(1,i) \times SR(1,i) \times [1 - q(i) \times qr(i) ]$
= probability of a response (incidence) in the first T years of life (i.e., up to the T-th birthday, age T ) at dose 0 (no exposure in addition to background exposure)

## Derivation of Formulas:

Background Risk of an Incidence:

$R_0 = \sum_{i=1,\ldots,T} \{ h(i) / [ h(i)^* + \mathbf{h(i)} ] \} \times S(1,i) \times \mathbf{SR(1,i)} \times [1 - q(i) \times \mathbf{qr(i)} ]$
= probability of a response (incidence) in the first T years of life
(i.e., up to the T-th birthday, age T ) at dose 0
(no exposure in addition to background exposure)

Contrast with the <u>form</u> of the calculation for the
Background Risk of a Mortality
and that h(i) refers to mortality here and incidence above:

$R_0 = \sum_{i=1,\ldots,T} [ h(i) / h(i)^* ] \times S(1,i) \times [1 - q(i) ]$
= probability of a response mortality in the first T years of
life (i.e., up to the T-th birthday, age T ) at dose 0
(no exposure in addition to background exposure)

## Derivation of Formulas: Risk with exposure

i=1, 2, ..., T

q(i) = exp [ - h(i)* ] = probability of surviving year i without exposure
when all causes of death are acting
conditional on the person surviving through year i-1

h(i)* = mortality rate due to all causes in year i without exposure
conditional on the person surviving through year i-1

h(i) = response (e.g., lung cancer) incidence rate in year i without exposure
conditional on the person not having the response through year i-1

qr(i) = exp [ - h(i) ] = probability of no response in year i without exposure
conditional on the person not responding through year i-1

f(i) = proportional effect (multipler) in year i assuming a proportional hazards
model for the effect of exposure of the form h(i) × f(i)
f(i) = [ 1 + e(i) ] if the multiplier is a linear function

h(i) × f(i) = response (e.g., lung cancer) incidence rate in year i with exposure
conditional on the person not having the response
through year i-1

## Derivation of Formulas: Risk with exposure

**i=1, 2, ..., T**

**q(i) = exp [ - h(i)* ] = probability of surviving year i without exposure when all causes of death are acting conditional on the person surviving through year i-1**

**h(i)* = mortality rate due to all causes in year i without exposure conditional on the person surviving through year i-1**

**h(i) = response (e.g., lung cancer) incidence rate in year i without exposure conditional on the person not having the response through year i-1**

**qr(i) = exp [ - h(i) ] = probability of no response (incidence) in year i without exposure conditional on the person not responding through year i-1**

**f(i) = proportional effect (multipler) in year i assuming a proportional hazards model for the effect of exposure of the form h(i) × f(i) f(i) = [ 1 + e(i) ] if the multiplier is a linear function**

**h(i) × f(i) = response (e.g., lung cancer) incidence rate in year i with exposure conditional on the person not having the response through year i-1**

A person is "observed" in year i if that person either dies in year i or has the specified response (incidence) in year i.

h(i)* + h(i) × f(i) = observation rate due to all causes in year i with exposure conditional on the person not dying or having the response through year i-1

## Derivation of Formulas: Risk with exposure

**i=1, 2, ..., T**

**q(i) = = exp[ - h(i)* ] = probability of surviving year i without exposure when all causes of death are acting conditional on the person surviving through year i-1**

**h(i)* = mortality rate due to all causes in year i without exposure conditional on the person surviving through year i-1**

**h(i) = response (e.g., lung cancer) incidence rate in year i without exposure conditional on the person not having the response through year i-1**

**qr(i) = exp[ -h(i) ] = probability of no response (incidence) in year i without exposure conditional on the person not responding through year i-1**

**f(i) = proportional effect (multipler) in year i assuming a proportional hazards model for the effect of exposure of the form h(i) × f(i) f(i) = [ 1 + e(i) ] if the multiplier is a linear function**

**h(i) × f(i) = response (e.g., lung cancer) incidence rate in year i with exposure conditional on the person not having the response through year i-1**

**h(i)* + h(i) × f(i) = observation rate due to all causes in year i with exposure conditional on the person not dying or having the response through year i-1**

exp { - h(i)* - h(i) × f(i) } = q(i) × exp { - h(i) × f(i) }
= q(i) × exp{ - h(i) - h(i) × [f(i) - 1] } = q(i) × qr(i) × exp {- h(i) × [ f(i)-1] }
probability with exposure of not dying and not responding in year i conditional on not dying and not responding thru year i-1

## Derivation of Formulas: Risk with exposure

i=1, 2, ..., T

q(i) = = exp[ - h(i)* ] = probability of surviving year i without exposure when all causes of death are acting conditional on the person surviving through year i-1

h(i)* = mortality rate due to all causes in year i without exposure conditional on the person surviving through year i-1

h(i) = response (e.g., lung cancer) incidence rate in year i without exposure conditional on the person not having the response through year i-1

qr(i) = exp[ -h(i) ] = probability of no response (incidence) in year i without exposure conditional on the person not responding through year i-1

S(1,i) = probability of surviving up to year i is the product of surviving each prior year: $S(1,i) = q(1) \times q(2) \times ... \times q(i-1)$ with S(1,1) = 1.0.

SR(1,i) = probability of no response up to year i is the product of no response in each prior year: $SR(1,i) = qr(1) \times qr(2) \times ... \times qr(i-1)$ with SR(1,1) = 1.0.

q(i) × qr(i) × exp {- h(i) × [ f(i)-1] } = probability with exposure of not dying and not responding in year i conditional on not dying and not responding thru year i-1

$q(1)\times qr(1)\times exp\{-h(1)\times[ f(1)-1 ]\} \times ... \times q(i-1)\times qr(i-1)\times exp\{-h(i-1)\times[ f(i-1)-1 ]\}$

$= S(1,i) \times SR(1,i) \times exp ( - \sum_{k=1,...,i-1} \{ - h(k) \times [ f(k) - 1 ] \} )$

= probability with exposure of not dying and not responding up to year i

$S(1,i)\times SR(1,i)\times exp(- \sum_{k=1,...,i-1} \{-h(k)\times[f(k)-1]\})\times[1-q(i)\times qr(i)\times exp\{-h(i)\times[ f(i)-1]\}]$

= probability with exposure of not dying and not responding up to year i and then dying (from any cause) or having the response in year i

## Derivation of Formulas: Risk with exposure

i=1, 2, ..., T

q(i) = = exp[ - h(i)* ] = probability of surviving year i without exposure when all causes of death are acting conditional on the person surviving through year i-1

h(i)* = mortality rate due to all causes in year i without exposure conditional on the person surviving through year i-1

h(i) = response (e.g., lung cancer) incidence rate in year i without exposure conditional on the person not having the response through year i-1

qr(i) = exp[ -h(i) ] = probability of no response (incidence) in year i without exposure conditional on the person not responding through year i-1

S(1,i) = probability of surviving up to year i is the product of surviving each prior year: $S(1,i) = q(1) \times q(2) \times ... \times q(i-1)$ with S(1,1) = 1.0.

SR(1,i) = probability of no response up to year i is the product of no response in each prior year: $SR(1,i) = qr(1) \times qr(2) \times ... \times qr(i-1)$ with SR(1,1) = 1.0.

q(i) × qr(i) × exp {- h(i) × [ f(i)-1] } = probability with exposure of not dying and not responding in year i conditional on not dying and not responding thru year i-1

$S(1,i) \times SR(1,i) \times exp(-\sum k=1,...,i-1 \{-h(k) \times[ f(k) - 1 ]\}) \times [1 - q(i) \times qr(i) \times exp\{-h(i)\times[ f(i)-1]\} ]$
= probability with exposure of not dying and not responding up to year i and then dying (from any cause) or having the response in year i

{ h(i) × f(i) } / { h(i)* + h(i) × f(i) } = proportion of observations (deaths plus incidences) in year i due to the response with exposure

$(\{h(i)\times f(i)\}/\{h(i)^*+h(i)\times f(i)\}) \times S(1,i) \times SR(1,i) \times exp(- \sum_{k=1,...,i-1}\{-h(k)\times[f(k)-1]\})$
$\times [1-q(i)\times qr(i)\times exp\{-h(i)\times[ f(i)-1]\}]$ = probability with exposure of not dying and not responding up to year i and then having the response in year i

## Derivation of Formulas: Risk with exposure

$$( \{ h(i) \times f(i) \} / \{ h(i)^* + h(i) \times f(i) \} )$$
$$\times S(1,i) \times SR(1,i) \times \exp(- \sum_{k=1,\ldots,i-1} \{-h(k) \times [f(k)-1]\})$$
$$\times [ 1 - q(i) \times qr(i) \times \exp \{ -h(i) \times [ f(i)-1] \} ]$$

= probability of not dying and not responding in i-1 years and then having the response in year i with exposure

$$R_{exposure} = \sum_{i=1,\ldots,T}$$
$$( \{h(i) \times f(i)\} / \{ h(i)^* + h(i) \times f(i) \} )$$
$$\times S(1,i) \times SR(1,i) \times \exp(- \sum_{k=1,\ldots,i-1} \{ -h(k) \times [ f(k)-1 ] \} )$$
$$\times [ 1 - q(i) \times qr(i) \times \exp \{ -h(i) \times [ f(i)-1] \} ]$$

= probability of a response (incidence) in the first T years of life (i.e., up to the T-th birthday, age T ) with exposure (with exposure in addition to the background exposure)

## Derivation of Formulas:

Risk of an Incidence with exposure:

$$R_{exposure} = \sum_{i=1,\ldots,T}$$
$$( \{h(i) \times f(i)\} / \{ h(i)^* + h(i) \times \mathbf{f(i)} \} )$$
$$\times S(1,i) \times \mathbf{SR(1,i)} \times \exp(- \sum_{k=1,\ldots,i-1} \{ -h(k) \times [ f(k)-1 ] \} )$$
$$\times [ 1 - q(i) \times \mathbf{qr(i)} \times \exp \{ -h(i) \times [ f(i)-1] \} ] )$$

Contrast with the <u>form</u> of the calculation for the Risk of a Mortality with exposure and that h(i) refers to mortality here and incidence above:

$$R_{exposure} = \sum_{i=1,\ldots,T} ( \{h(i) \times f(i)\} / \{h(i)^* + h(i) \times [ f(i)-1 ] \} )$$
$$\times S(1,i) \times \exp(- \sum_{k=1,\ldots,i-1} \{ -h(k) \times [ f(k)-1 ] \} )$$
$$\times (1 - q(i) \times \exp \{ - h(i) \times [ f(i) - 1 ] \} )$$

## Risks

$R_0 = \sum_{i=1,\ldots,T} \{ h(i) / [ h(i)^* + h(i) ] \} \times S(1,i) \times SR(1,i) \times [1 - q(i) \times qr(i) ]$
= probability of a response (incidence) in the first T years of life (i.e., up to the T-th birthday, age T )
at dose 0 (no exposure in addition to background exposure)

$R_{exposure} = \sum_{i=1,\ldots,T}$

$( \{h(i) \times f(i)\} / \{ h(i)^* + h(i) \times f(i) \} )$
$\times S(1,i) \times SR(1,i) \times \exp(- \sum_{k=1,\ldots,i-1} \{ -h(k) \times [ f(k)-1 ] \} )$
$\times [ 1 - q(i) \times qr(i) \times \exp \{ -h(i) \times [ f(i)-1] \} ]$

= probability of a response (incidence) in the first T years of life (i.e., up to the T-th birthday, age T ) with exposure (with exposure in addition to the background exposure)

Added Risk = $R_{exposure} - R_0$

Extra Risk = $( R_{exposure} - R_0 ) / ( 1 - R_0 )$

Excess Risk = either Added Risk or Extra Risk